(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 855 805 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
29.07.1998 Bulletin 1998/31

(51) Int. Cl.[6]: H04B 1/66

(21) Application number: 97310098.5

(22) Date of filing: 15.12.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 22.01.1997 JP 9605/97

(71) Applicant:
SHARP KABUSHIKI KAISHA
Osaka-shi, Osaka-fu 545 (JP)

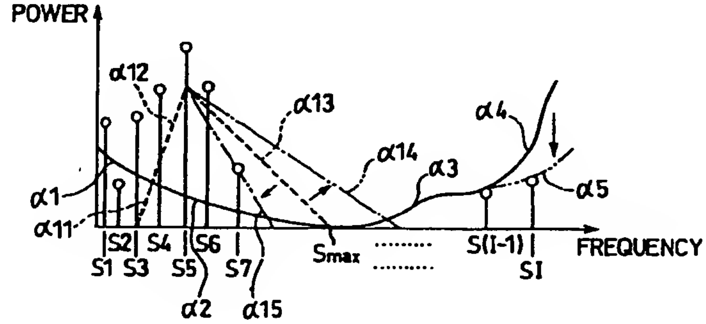
(72) Inventor: Fujii, Osamu
Higashihiroshima-shi, Hiroshima 739-01 (JP)

(74) Representative:
Jenkins, Richard Gray et al
R.G.C. Jenkins & Co.
26 Caxton Street
London SW1H 0RJ (GB)

(54) Method of encoding digital audio signals

(57) The method of encoding digital data in the present invention enables change of minimum limit of audibility characteristics and/or masking characteristics, which are usually set on the basis of the aural-psychological characteristics of persons with typical hearing, thus changing the allocation of quantized bits to each frequency band and allowing selection of a sound quality which accords with the listener's hearing. The present invention is suitable for ATRAC, a method for compressed encoding for mini-discs.

FIG.1



EP 0 855 805 A2

wave signal, the number of bits allocated to the frequency bands f1 which include the sine wave signal becomes relatively smaller, the quantization error of the sine wave signal becomes greater, and sound quality deteriorates.

In regard to this point, the present Applicant has proposed, in Japanese Unexamined Patent Publication 7-202823/1995, a structure which automatically limits the number of bits which may be allocated to frequency bands with low power S. However, a drawback of this conventional art is that, since the maximum number of bits which may be allocated to each frequency band is determined on the basis of its power, when the power of white noise is large, there are cases when no limitation on bit allocation to that frequency band is made.

SUMMARY OF THE INVENTION

One object of the present invention is to provide a method of encoding digital data capable of attaining a sound quality which accords with the listener's hearing.

Another object of the present invention is to provide a method of encoding digital data capable of preventing deterioration of sound quality even of signals with narrow spectrum bands.

In order to realize the first object mentioned above, the first method of encoding digital data of the present invention encodes digital data such as musical tones and sounds by converting it into frequency domains, dividing the converted spectra into a plurality of frequency bands, changing a minimum limit of audibility characteristic so as to set a masking threshold, and allocating quantized bits for each frequency band in accordance with ratios of masking threshold to noise which are found for each frequency band in accordance with power or energy of each frequency band in consideration of aural-psychological characteristics.

The above structure, by enabling change of the minimum limit of audibility characteristic among aural-psychological characteristics, frees aural-psychological characteristics from definition by the characteristics of persons with typical hearing, and makes possible selection of whether or not to allocate bits to spectra with small inaudible domains, or spectra with ultra-low or ultra-high domains. Accordingly, it becomes possible to respond to persons with superior hearing or to individual, subjective preference, and sound quality which accords with listeners' hearing can be attained.

Next, in order to realize the first object mentioned above, the second method of encoding digital data of the present invention encodes digital data such as musical tones and sounds by converting it into frequency domains, dividing the converted spectra into a plurality of frequency bands, changing a masking characteristic so as to set a masking threshold, and allocating quantized bits for each frequency band in accordance with ratios of the masking threshold to noise for each frequency band which are found in accordance with power or energy of each frequency band in consideration of aural-psychological characteristics.

The above structure, by enabling change of the masking characteristic among the aural-psychological characteristics, frees aural-psychological characteristics from definition by the characteristics of persons with typical hearing, and makes possible selection of whether to allocate bits, for example, to spectra which, for example, suffer masking in a critical band. Accordingly, it becomes possible to respond to persons with superior hearing or to individual, subjective preference, and sound quality which accords with listeners' hearing can be attained.

Next, in order to realize the first object mentioned above, the third method of encoding digital data of the present invention encodes digital data such as musical tones and sounds by converting it into frequency domains, dividing the converted spectra into a plurality of frequency bands, and switching among (i) bit allocation in accordance with ratios of masking threshold to noise which are found for each frequency band in accordance with power or energy of each frequency band in consideration of aural-psychological characteristics, (ii) bit allocation in accordance with a representative value of the power or the energy of each frequency band, and (iii) bit allocation giving weight to each of the foregoing bit allocation methods.

With respect to data, such as white noise having a spectral composition which is comparatively flat, the above structure makes possible bit allocation which is flat along the frequency axis. Again, with respect to data, such as sine wave signals, with narrow band width, the above structure makes possible bit allocation which emphasizes the signal with narrow band width. Accordingly, selection of a sound quality which is suited to the source of the musical tone is made possible.

Finally, the fourth method of encoding digital data of the present invention, in order to realize the second object mentioned above, switches among bit allocation methods (i), (ii), and (iii) described in the third method of encoding digital data in accordance with a relationship between the masking threshold and peaks and local peaks found based on differences in power or energy between adjacent spectra within each frequency band.

The above structure makes it possible to automatically allocate bits according to the method most suited to the digital data, whether it is white noise or other data with wide band width, or sine wave signals or other data with narrow band width, thus preventing deterioration of sound quality, even with musical tones not suited to bit allocation using simultaneous masking such as the masking threshold/noise ratio.

The other objects, features, and superior points of the present invention will be made clear by the description below. Further, the advantages of this invention will be evident from the following explanation in reference to the Figures.

control microcomputer 36. In association with the system control microcomputer 36 is provided an input operation means, which enables sound-quality selection operations, which will be discussed below, as well as song title input, song selection operations, etc.

Next, the bit allocation method in the first embodiment of the present invention, which is performed according to the ATRAC method by the audio compression circuit 6 of the mini-disc recording and reproduction device 1 structured as described above, will be explained, referring to Figures 1 and 3.

In the ATRAC method, the audio data sampled at 44.1 kHz, as mentioned above, is divided into certain frequency bands, specifically a Low frequency band from 0 kHz to 5.5 kHz, a Middle frequency band from 5.5 kHz to 11 kHz, and a High frequency band from 11 kHz to 22 kHz, and the audio data bridging certain time frames for each divided frequency band is converted, by means of the MDCT processing, into an MDCT coefficient, which is the data of one frequency domain. The MDCT coefficients converted in this manner are then converted into spectrum powers Si for i number of frequency bands (i = 1, 2, ..., I, with I equal to, for example, 25). Processing like that shown in Figure 3 is then carried out to allocate quantized bits in accordance with each spectrum power Si thus obtained.

The audio compression circuit 6 includes a table ROM 6a, and in the table ROM 6a are stored masking characteristics and/or minimum limit of audibility characteristics according to the ATRAC method. These minimum limit of audibility characteristics appear as a curve shown by reference symbols $\alpha1$, $\alpha2$, $\alpha3$, and $\alpha4$ on Figure 1. The masking characteristics, calculated in accordance with the spectrum powers Si, a critical band width of each frequency band, etc., appear, for a power distribution like that shown in Figure 1, for example, as a curve shown by reference symbols $\alpha11$, $\alpha12$, and $\alpha13$. The minimum limit of audibility characteristics shown by the reference symbols $\alpha1$ through $\alpha4$ and masking characteristics shown by reference symbols $\alpha11$ through $\alpha13$ are prepared in accordance with the aural-psychological characteristics of persons with typical hearing characteristics, and are fixed characteristics.

However, in the first embodiment of the present invention, the minimum limit of audibility and/or the masking characteristics can be changed. In concrete terms, for example in the case of the masking characteristics, the greater the spectrum power and the higher the frequency, the larger the range of masking of other frequency bands. In the example in Figure 1, the maximum limit Smax of the range influenced by spectrum power S5, which is a peak power, is shown by $\alpha13 \times (1 \pm \Sigma k)$. Here, Σk is a coefficient for weighting. If a plurality of variables k are stored in advance in the table ROM 6a, and the variables k are switched by means of a register 36a in the system control microcomputer 36, the masking characteristic curve $\alpha13$ can be changed within the range from $\alpha14$ through $\alpha15$. The variable k can be set by the listener through the input operating means 37.

For example, by changing the masking characteristic curve from $\alpha13$ to $\alpha14$, the band masked is widened, the level of masking is increased, and the number of bits allocated to signals with low power is decreased, or even eliminated. Accordingly, bit allocation to signals of relatively greater power is increased, and the dynamic range of the high-power signals is increased. If, on the other hand, the masking characteristic curve is changed from $\alpha13$ to $\alpha15$, bit allocation to low-power signals is increased, and bit allocation to signals of relatively greater power is decreased. Accordingly, the frequency range can be enlarged. The same effect can also be obtained by giving the masking characteristic curve $\alpha13$ an offset instead of weighting.

In the same way, with regard to the minimum limit of audibility characteristics, the minimum limit of audibility characteristic curve $\alpha1$ through $\alpha4$, which is based on the aural-psychological characteristics of persons with typical hearing characteristics, can be weighted or given an offset, thereby changing the $\alpha4$ portion of the curve, for example, as shown by reference symbol $\alpha5$. In this way, relatively more bits are allocated to the high-frequency bands.

Next, processing for allocation of quantized bits will be explained, referring to Figure 3. First, in Step p1, the spectrum power Si of each frequency band is calculated from the sum of squares of the MDCT coefficients for that frequency band (which are obtained by means of the MDCT processing). In Step p2, the audio compression circuit 6 selects, through the register 36a of the system control microcomputer 36, parameters for change of masking characteristics, such as the variables k, which are stored in the table ROM 6a. In Step p3, in the same way as in Step p2, parameters for change of the minimum limit of audibility characteristics are selected.

In Step p4, reference masking characteristics and minimum limit of audibility characteristics previously calculated and stored in the table ROM 6a are changed in accordance with the parameters selected in Steps p2 and p3, and these two characteristics are synthesized in order to determine a final masking threshold. In other words, if the minimum limit of audibility characteristic curve thus changed is as shown by the reference symbols $\alpha1$, $\alpha2$, $\alpha3$, $\alpha5$, and the masking characteristic curve thus changed is as shown by the reference symbols $\alpha11$, $\alpha12$, $\alpha14$, the curve of the final masking threshold obtained by synthesis will be as shown by the reference symbols $\alpha1$, $\alpha12$, $\alpha14$, $\alpha3$, $\alpha5$.

In Step p5, if the index of each frequency band is i, the ratio of the frequency band's spectrum power Si (calculated in Step p1) to its masking threshold Mi (calculated in Step p4) SMRi = Si/Mi is calculated for all frequency bands. On a logarithmic graph, the ratio SMRi for each frequency band will correspond to that

cases where it resembles a single sine wave, for example with a solo piano piece, if the bit allocation is performed only according to the ratio of masking threshold to quantized noise MNRi(n), many bits will be allocated to noise elements with low power, and the error in quantizing of the piano becomes relatively great. However, if the bit allocation percentage x can be changed as outlined above, the bit allocation according to the power of quantized noise SNi(n) is carried out in addition to that according to the ratio of masking threshold to quantized noise MNRi(n), thereby ensuring that the number of bits allocated to the piano can be increased, the error in quantizing of the piano is reduced.

Again, if the input signal is composed of sound with many local peaks and noise, for example an orchestra piece, the bit allocation can be performed in accordance with the ratio of masking threshold to quantized noise MNRi(n), in which the noise and the musical tones composing small local peaks in bands close to large signals can be masked, thus allocating no bits to them, and more bits can be allocated to large signals which are not masked. This enables high fidelity recording.

Further, with input signals lying between the foregoing two examples, which are composed of a musical tone with three or four local peaks and noise, for example a solo clarinet piece, by giving weight both to the bit allocation according to the ratio of masking threshold to quantized noise MNRi(n) and to the bit allocation according to the power of quantized noise SNi(n), fidelity of the clarinet can be improved.

In this way, the bit allocation method most suited to any musical tone source can be selected.

The third embodiment of the present invention will next be explained, in reference to Figures 5 and 6.

Figure 5 is a flow chart for explaining the bit allocation method in the third embodiment of the present invention. The notable feature of this bit allocation method is that the percentage x of (a) the bit allocation according to the ratio of masking threshold to quantized noise MNRi(n) to (b) the bit allocation according to the power of quantized noise SNi(n) is automatically determined on the basis of the relationship between (1) peaks and local peaks in spectrum powers Si and (2) masking thresholds.

First, the peak value among the spectrum powers of all frequency bands from S1 to Sl, such as that shown by reference symbol S5 on Figure 6, is found. Then a masking threshold, such as that shown on Figure 6, which includes masking characteristics due to that peak level, is found. Next, local peaks such as that shown by reference symbol S8 on Figure 6 are found for each frequency band. The number of such local peaks masked by the masking threshold, and the number of such local peaks not so masked, are respectively found, and the ratio between masked local peaks and unmasked local peaks determines the percentage x.

In other words, if the total number of local peaks is NM, and the number of masked local peaks is M, then:

$$M/(NM+1) = 0 \tag{1}$$

Accordingly, if there are no masked local peaks, the percentage x will be 0%, and the number of bits available for the first allocation B1 will be set at 0. If, on the other hand,

$$0 < M/(NM+1) < 0.5 \tag{2}$$

then x is from 50% to 90%, and if

$$0.5 < M/(NM+1) \tag{3}$$

then x is 100%, and the number of bits available for the first allocation B1 will be the entirety of total available bits B0.

Here the detection of local peaks and the selection of the percentage x will be discussed. The local peaks are found for all the frequency bands after the peak spectrum power (S5 in Figure 6) is found. In the example in Figure 6, the difference D34, D45, ..., D89 and its polarity between each spectrum power S3 to S9 within a certain number of frequency bands from peak value S5 (in the example in Figure 6, two frequency bands on the low-frequency side and four on the high-frequency side) is found, and the local peaks are detected on the basis of change in polarity and the absolute value of those differences. In this way, the local peaks are found across all frequency bands. In concrete terms, in the case of Figure 6, there is only one local peak (S8), and that local peak is masked by the masking threshold; thus $M/(NM+1) = 1/(1+1) = 0.5$. Accordingly, equation (2) above will be applied, and a percentage x = 50% to 90% will be selected.

Next, the bit allocation method of the third embodiment will be explained in reference to Figure 5.

In this bit allocation method, after the spectrum power Si of each frequency band has been calculated in Step p21 as in Step p11 and Step p1 above, the peak value is found in Step p22, and the masking threshold including the masking characteristics of that peak value is found in Step p23. In Step p24, the percentage x is calculated by means of equations (1) through (3) above, and the number of bits available for the first allocation B1 is calculated. Then, in Steps p25 through p27, as in Steps p15 through p17 above, the first bit allocation, according to the ratio of masking threshold to quantized noise MNRi(n), is performed, and then in Steps p28 and p29, as in Steps p18 and p19 above, the second bit allocation, according to the power of quantized noise SNi(n), is performed.

In this way, a bit allocation with high sound quality appropriate to the musical tones like that shown in Figure 4 can be performed automatically, and deterioration of sound quality, even with respect to musical tones not suited to the bit allocation according to the ratio of masking threshold to quantized noise MNRi(n), can be prevented.

each frequency band;

wherein said steps (i) and (ii) are performed in accordance with the percentage so as to allocate the total number of quantized bits, thereby allocating the number of the quantized bits of each frequency band.

8. The method of encoding digital data according to Claim 7, wherein:

the percentage is determined in accordance with a relationship between the masking threshold and peaks and local peaks found based on differences in power or energy between adjacent spectra within each frequency band.

9. The method of encoding digital data according to Claim 7, wherein:

the percentage corresponds, if NM is the total number of the local peaks, with a ratio of the number M of the local peaks which are masked by the masking threshold to the number N of the local peaks which are not masked by the masking threshold.

10. The method of encoding digital data according to Claim 9, wherein:

the percentage is set to 0% when $M/(NM+1) = 0$ is satisfied; and
the percentage is set to 100% when $0.5 < M/(NM+1)$ is satisfied.

11. A method of encoding digital data in which, in converting digital data such as musical tones and sounds into frequency domains, dividing the converted spectra into a plurality of frequency bands, and allocating bits for each frequency band so as to encode the digital data, the allocating of bits is performed in accordance with a ratio of masking threshold to noise of each frequency band in consideration of aural-psychological characteristics, the ratio being calculated for each frequency band in accordance with power or energy of each frequency band, wherein:

minimum limit of audibility characteristics that have been previously found are changeable.

12. A method of encoding digital data in which, in converting digital data such as musical tones and sounds into frequency domains, dividing the converted spectra into a plurality of frequency bands, and allocating bits for each frequency band so as to encode the digital data, the allocating of bits is performed in accordance with a ratio of masking threshold to noise of each frequency band in consideration of aural-psychological characteristics, the ratio being calculated for each frequency band in accordance with power or energy of each frequency band, wherein:

masking characteristics that have been previously found in accordance with the power or the energy are changeable.

13. A method of encoding digital in which digital data such as musical tones and sounds is converted into frequency domains, the converted spectra are divided into a plurality of frequency bands, and bits are allocated for each frequency band so as to perform the encoding, said method comprising the steps of:

(i) performing a first allocation of the quantized bits in accordance with ratios of masking threshold to noise which are found for each frequency band in accordance with power or energy of each frequency band;
(ii) performing a second allocation of the quantized bits in accordance with a representative value of the power or the energy of each frequency band; and
(iii) performing a third allocation of the quantized bits giving weight to the bit allocation methods of each of said steps (i) and (ii);
wherein said steps (i), (ii), and (iii) are switchable.

14. The method of encoding digital data according to Claim 13, wherein:

said steps (i), (ii), and (iii) are switched in accordance with a relationship between the masking threshold and peaks and local peaks found based on differences in power or energy between adjacent spectra within each frequency band.

F I G. 2

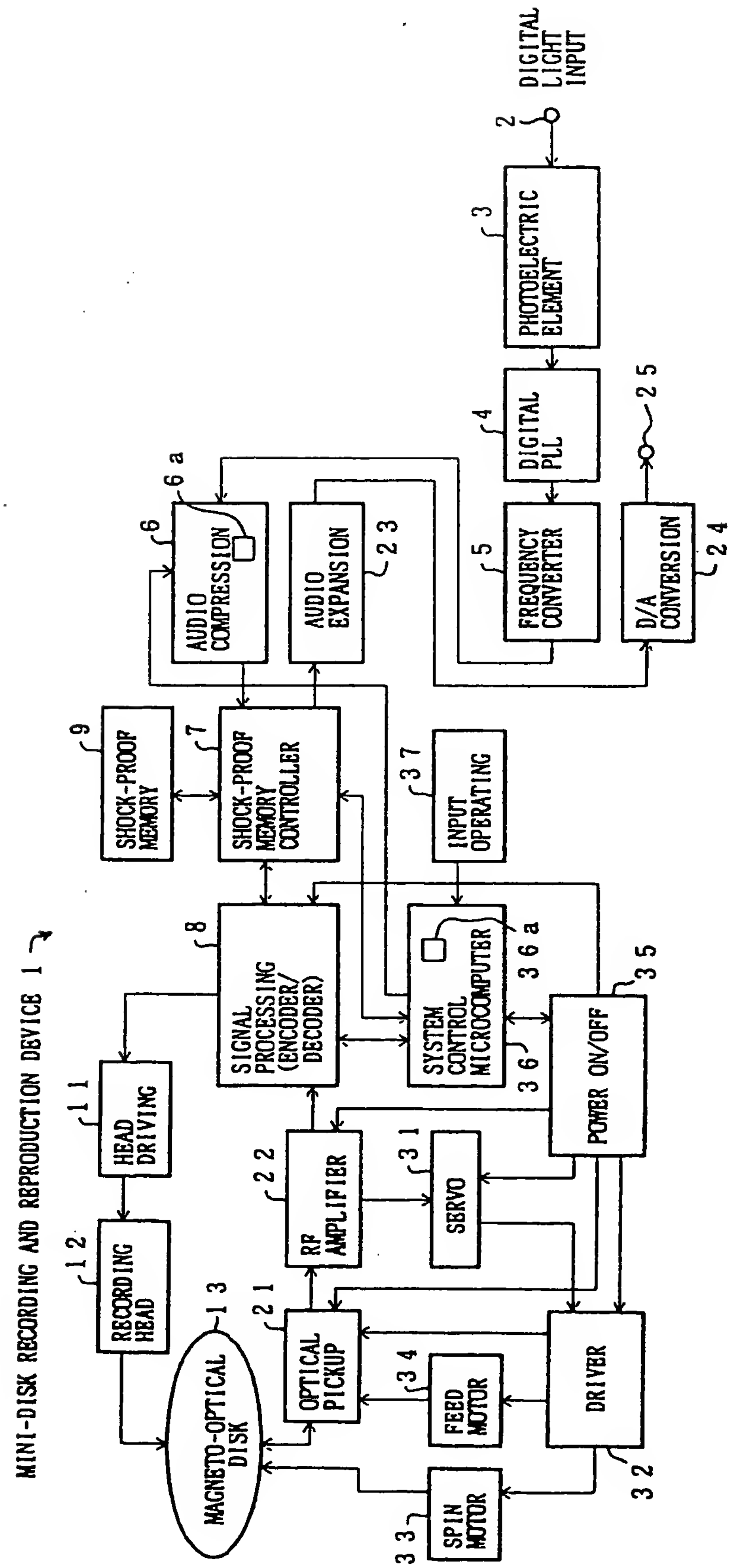
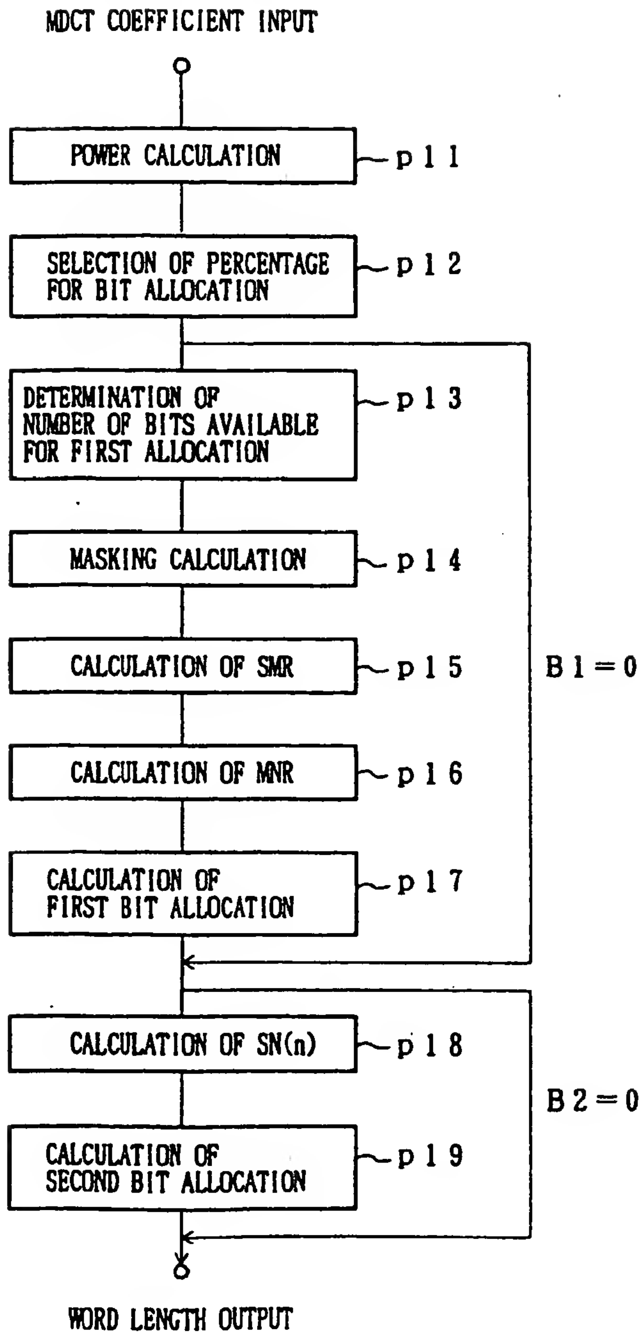

MINI-DISK RECORDING AND REPRODUCTION DEVICE 1
DIGITAL LIGHT INPUT
2
3
PHOTOELECTRIC ELEMENT
4
DIGITAL PLL
5
FREQUENCY CONVERTER
25
24
D/A CONVERSION
6a
6
AUDIO COMPRESSION
23
AUDIO EXPANSION
9
SHOCK-PROOF MEMORY
7
SHOCK-PROOF MEMORY CONTROLLER
37
INPUT OPERATING
8
SIGNAL PROCESSING (ENCODER/DECODER)
36a
36
SYSTEM CONTROL MICROCOMPUTER
35
POWER ON/OFF
11
HEAD DRIVING
12
RECORDING HEAD
13
MAGNETO-OPTICAL DISK
22
RF AMPLIFIER
31
SERVO
21
OPTICAL PICKUP
34
FEED MOTOR
33
SPIN MOTOR
32
DRIVER

FIG. 4


MDCT COEFFICIENT INPUT
POWER CALCULATION
p11
SELECTION OF PERCENTAGE FOR BIT ALLOCATION
p12
DETERMINATION OF NUMBER OF BITS AVAILABLE FOR FIRST ALLOCATION
p13
MASKING CALCULATION
p14
CALCULATION OF SMR
p15
B1=0
CALCULATION OF MNR
p16
CALCULATION OF FIRST BIT ALLOCATION
p17
CALCULATION OF SN(n)
p18
B2=0
CALCULATION OF SECOND BIT ALLOCATION
p19
WORD LENGTH OUTPUT

F I G. 5

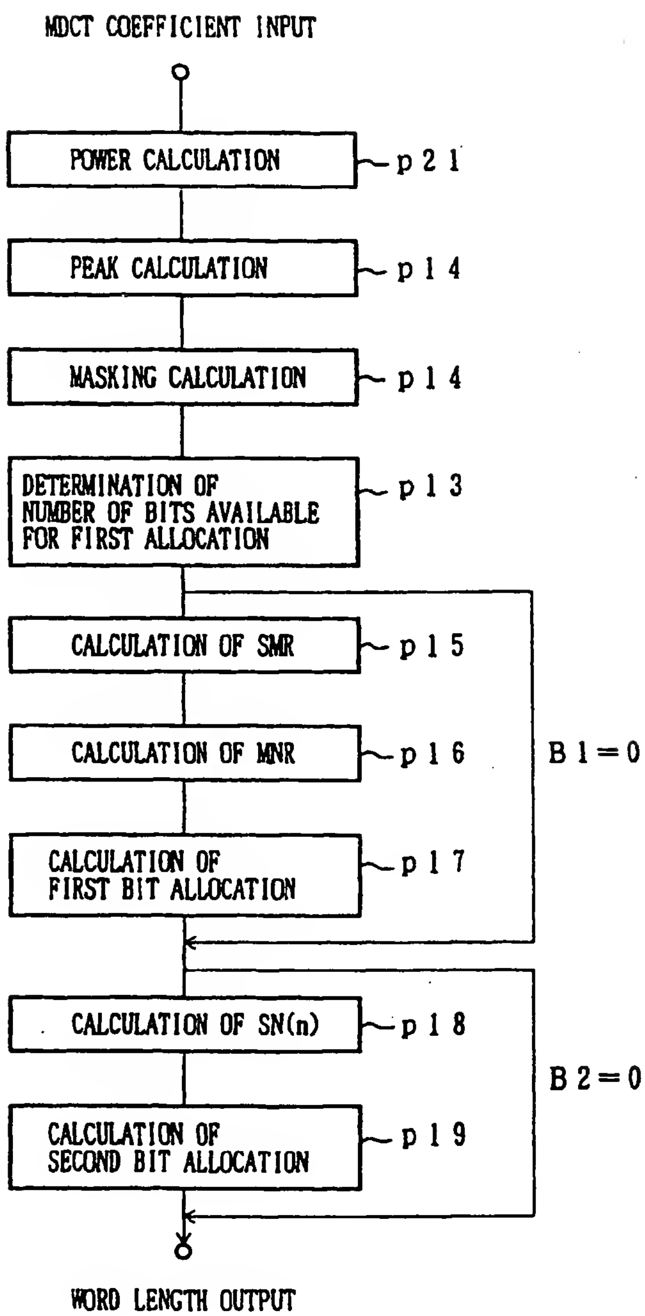
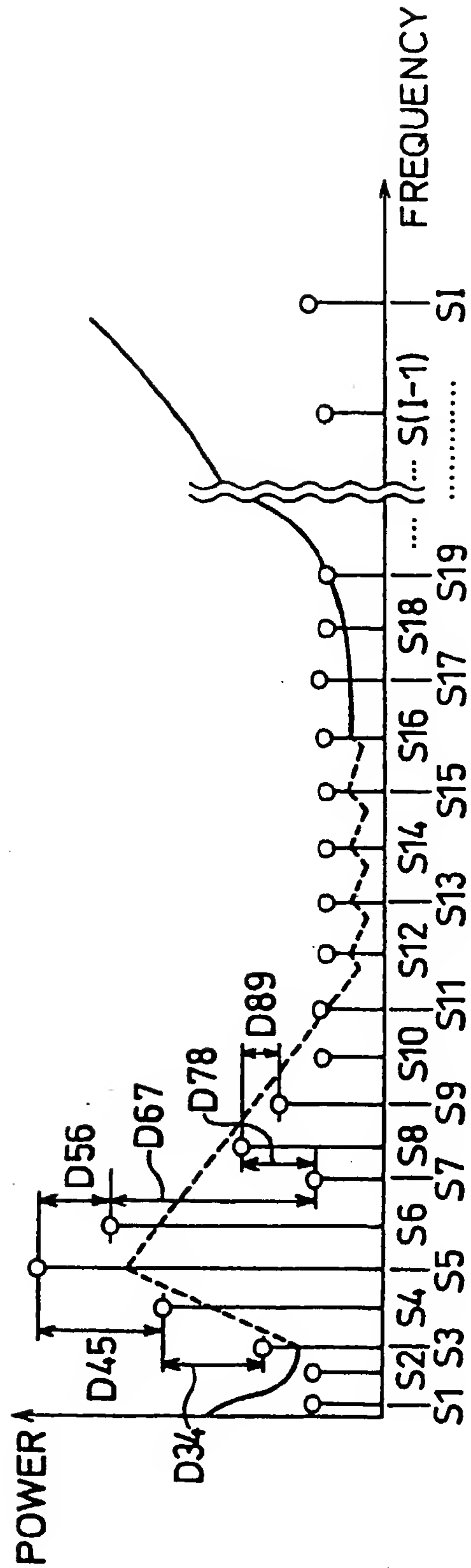

MDCT COEFFICIENT INPUT
POWER CALCULATION
p21
PEAK CALCULATION
p14
MASKING CALCULATION
p14
DETERMINATION OF NUMBER OF BITS AVAILABLE FOR FIRST ALLOCATION
p13
CALCULATION OF SMR
p15
CALCULATION OF MNR
p16
CALCULATION OF FIRST BIT ALLOCATION
p17
B1=0
CALCULATION OF SN(n)
p18
CALCULATION OF SECOND BIT ALLOCATION
p19
B2=0
WORD LENGTH OUTPUT

FIG. 6


POWER
FREQUENCY
D34
D45
D56
D67
D78
D89
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
S13
S14
S15
S16
S17
S18
S19
S(I-1)
SI

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 855 805 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
19.04.2000 Bulletin 2000/16

(51) Int. Cl.[7]: H04B 1/66

(43) Date of publication A2:
29.07.1998 Bulletin 1998/31

(21) Application number: 97310098.5

(22) Date of filing: 15.12.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 22.01.1997 JP 960597

(71) Applicant:
SHARP KABUSHIKI KAISHA
Osaka-shi, Osaka-fu 545-0013 (JP)

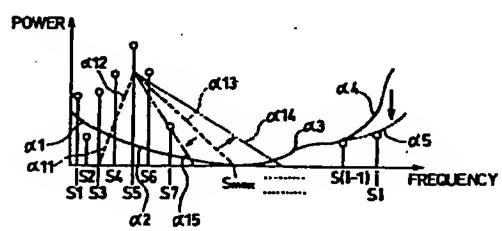
(72) Inventor: Fujii, Osamu
Higashihiroshima-shi, Hiroshima 739-01 (JP)

(74) Representative:
Abbie, Andrew Kenneth et al
R.G.C. Jenkins & Co.
26 Caxton Street
London SW1H 0RJ (GB)

(54) Method of encoding digital audio signals

(57) The method of encoding digital data in the present invention enables change of minimum limit of audibility characteristics and/or masking characteristics, which are usually set on the basis of the aural-psychological characteristics of persons with typical hearing, thus changing the allocation of quantized bits to each frequency band and allowing selection of a sound quality which accords with the listener's hearing. The present invention is suitable for ATRAC, a method for compressed encoding for mini-discs.

FIG.1



EP 0 855 805 A3

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 97 31 0098

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

29-02-2000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0661821 A	05-07-1995	JP 7202823 A	04-08-1995
		DE 69418994 D	15-07-1999
		DE 69418994 T	02-12-1995
		US 5684922 A	04-11-1997
EP 0570131 A	18-11-1993	JP 5313694 A	26-11-1993
		DE 69325312 D	22-07-1999
		DE 69325312 T	16-12-1999
		US 5553193 A	03-09-1996
EP 0612160 A	24-08-1994	JP 6242798 A	02-09-1994

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82